

The Möbius function and statistical mechanics

F. Cellarosi · Ya. G. Sinai

Received: 10 December 2010 / Accepted: 28 April 2011 / Published online: 15 June 2011
© The Author(s) 2011. This article is published with open access at SpringerLink.com

Abstract We consider a probabilistic model for square-free numbers, and provide limit theorems for several random variables defined in our ensemble. The limit transition corresponds to the thermodynamical limit in Statistical Mechanics. We also prove some inequalities inspired by a recent conjecture by P. Sarnak concerning the randomness in the Möbius sequence, and discuss a method of summation for the Riemann zeta function $\zeta(s)$ on the vertical line $\Re s = 1$.

1 Introduction

The purpose of this paper is to study the statistical properties of the classical Möbius function from Number Theory. It is defined for positive integers n by the formula

$$\mu(n) = \begin{cases} 1 & \text{if } n=1; \\ 0 & \text{if } n \text{ is not square-free;} \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct primes} \end{cases}$$

Communicated by A. Laptev.

F. Cellarosi (✉) · Ya. G. Sinai
Mathematics Department, Princeton University,
Princeton, NJ, USA
e-mail: fcellaro@math.princeton.edu

Ya. G. Sinai
Landau Institute of Theoretical Physics, Russian Academy of Sciences,
Moscow, Russia
e-mail: sinai@math.princeton.edu

and is closely connected with the Riemann ζ -function by the formulae

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}, \quad \sum_{n=1}^{\infty} \frac{\mu^2(n)}{n^s} = \frac{\zeta(s)}{\zeta(2s)}. \quad (1)$$

In [31], Sarnak raised a general question: to what extent can the sequence $\{\mu(n)\}_{n \in \mathbb{N}}$ be treated as a typical realization of some ergodic random process. A possible approach to this problem requires the proof of the so-called *Chowla conjecture* about the existence of the limits giving arbitrary finite-dimensional distributions of this process. The proof of this hypothesis is presumably very difficult.

Below we study the randomness of μ using some ideas from Statistical Mechanics. Fix m and introduce the ensemble Ω_m , consisting of integers having the form $p_1^{\nu_1} p_2^{\nu_2} \cdots p_m^{\nu_m}$, where $2 = p_1 < p_2 < \cdots < p_m$ are the first m prime numbers and $\nu_j \in \{0, 1\}$. It is clear that $\mu(n) \neq 0$ for every $n \in \Omega_m$. Moreover, every square free number $n \leq p_m$ belongs to Ω_m . As n grows above p_m , the set of $n \in \Omega_m$ becomes much smaller than the set of all n for which $\mu(n) \neq 0$. The cardinality of Ω_m is clearly 2^m and its largest element is $p_1 p_2 \cdots p_m$. Notice that $\Omega_m \subseteq \Omega_{m+1}$ and the union $\bigcup_{m \geq 1} \Omega_m$ is the set of all square-free numbers.

Introduce the probability distribution Π_m on Ω_m for which

$$\pi_m(n) = \frac{1}{Z_m n}, \quad n \in \Omega_m, \quad (2)$$

where Z_m is the normalizing factor such that $\sum_{n \in \Omega_m} \pi_m(n) = 1$. By analogy with Statistical Mechanics, the probability distribution Π_m can be called *microcanonical distribution* and Z_m *partition function*. It will be shown that Π_m plays an important role in our analysis.

The asymptotic expression for Z_m follows from the classical Mertens product formula [28]

$$\lim_{n \rightarrow \infty} \ln n \prod_{p \leq n} \left(1 - \frac{1}{p}\right) = e^{-\gamma} \approx 0.561459,$$

where γ is the Euler–Mascheroni constant, $\gamma = \lim_{n \rightarrow \infty} (\sum_{k=1}^n \frac{1}{k} - \ln n) \approx 0.57721$. In fact

$$\frac{1}{\ln n} \prod_{p \leq n} \left(1 + \frac{1}{p}\right) = \frac{\left(\prod_{p \leq n} \frac{1}{1-p^{-2}}\right)^{-1}}{\ln n \prod_{p \leq n} \left(1 - \frac{1}{p}\right)} \rightarrow \frac{\zeta(2)^{-1}}{e^{-\gamma}} \quad \text{as } n \rightarrow \infty.$$

Thus, as $m \rightarrow \infty$,

$$Z_m \sim \frac{e^{\gamma}}{\zeta(2)} \ln p_m \quad (3)$$

As in Statistical Mechanics, $Z_m \rightarrow \infty$ as $m \rightarrow \infty$ and this opens some possibilities for a thermodynamic limit transition.

In the ensemble (Ω_m, Π_m) the random variables v_j are mutually independent and

$$\Pi_m\{v_j = 0\} = \frac{p_j}{1 + p_j} = 1 - \frac{1}{1 + p_j}, \quad \Pi_m\{v_j = 1\} = \frac{1}{1 + p_j}, \quad 1 \leq j \leq m.$$

The proof of this fact follows easily from the representation

$$\pi_m(n) = \frac{1}{Z_m \prod_{j=1}^m p_j^{v_j}}.$$

It also implies that the probability distributions Π_m are compatible for different m , i.e. the restriction of Π_{m+1} to Ω_m gives Π_m . In this sense one can talk about the limiting objects Ω_∞ and Π_∞ .

Consider the entropy of the distribution Π_m . We have

$$\begin{aligned} H(\Pi_m) &= - \sum_{n \in \Omega_m} \pi_m(n) \ln \pi_m(n) = \ln Z_m + \sum_{j=1}^m \mathbb{E}_{\Pi_m}(v_j \ln p_j) \\ &= \ln Z_m + \sum_{j=1}^m \frac{\ln p_j}{1 + p_j} = \ln Z_m + \sum_{t=1}^{p_m} \frac{\ln t}{1 + t} (N(t) - N(t-1)) \\ &= \ln Z_m + \frac{N(p_m) \ln(p_m + 1)}{p_m + 2} - \sum_{t=1}^{p_m} N(t) \left(\frac{\ln(t+1)}{t+2} - \frac{\ln t}{t+1} \right) \\ &= O(1) + \ln Z_m - \sum_{t=2}^{p_m} \left(\frac{t}{\ln t} + O\left(\frac{t}{\ln^2 t}\right) \right) \left(\frac{1 - \ln t}{t^2} + O\left(\frac{1}{t^3}\right) \right) \\ &= O(1) + \ln Z_m + \sum_{t=2}^{p_m} \frac{1}{t} - \sum_{t=2}^{p_m} \frac{1}{t \ln t} = O(\ln \ln p_m) + \sum_{t=2}^{p_m} \frac{1}{t} \sim \ln p_m \end{aligned}$$

as $m \rightarrow \infty$. Even though $H(\Pi_m) \rightarrow \infty$ as $m \rightarrow \infty$, we have

$$\frac{-\ln \pi_m(n)}{H(\Pi_m)} = \frac{\ln Z_m + \sum_{j=1}^m v_j(n) \ln p_j}{H(\Pi_m)} \sim \sum_{j=1}^m v_j(n) \frac{\ln p_j}{\ln p_m}.$$

In Sect. 2, it will be shown that $\sum_{j=1}^m v_j(n) \frac{\ln p_j}{\ln p_m}$ does not have a limit in any sense, but it has a limiting distribution w.r.t. Π_m as $m \rightarrow \infty$. Therefore, the classical Shannon–McMillan theorem from Information Theory is not valid. This means that our case is different from the usual situation in Statistical Mechanics.

Introduce the random variables

$$\zeta_m = \sum_{j=1}^m v_j \frac{\ln p_j}{\ln p_m}.$$

In Sect. 2 we prove the following

Theorem 1 *As $m \rightarrow \infty$, the distributions of the random variables $\{\zeta_m\}_{m \in \mathbb{N}}$ converge weakly to a limiting probability distribution whose characteristic function $\varphi(\lambda)$ has the form*

$$\varphi(\lambda) = \exp \left\{ \int_0^1 \frac{e^{i\lambda v} - 1}{v} dv \right\}. \quad (4)$$

Moreover, the limiting probability distribution is infinitely divisible.

The corresponding probability distribution is called *Dickman–De Bruijn distribution* (see [9, 6]). It appears naturally in Probability Theory in the following way. Let $\{\eta_j\}$ be a sequence of independent random variables such that

$$P\{\eta_j = j\} = \frac{1}{j}, \quad P\{\eta_j = 0\} = 1 - \frac{1}{j} \quad (5)$$

and $\zeta_m = \sum_{j=1}^m \eta_j$. Then the limiting distribution of $\frac{1}{m}\zeta_m$ as $m \rightarrow \infty$ is the Dickman–De Bruijn distribution.

Let us recall that a random variable ξ called *infinitely divisible* if for every N , one can find N independent, identically distributed random variables ξ'_1, \dots, ξ'_N , such that $\xi = \xi'_1 + \dots + \xi'_N$. A complete characterization of such probability distributions in terms of the logarithm of their characteristic functions was provided by A.N. Kolmogorov, P. Lévy, and A.Ya. Khintchin, see, e.g. [14].

The density $d(t)$ of the distribution given by (4) is constant on the interval $[0, 1]$ where it takes the value $e^{-\gamma}$.

In particular, Theorem 1 and (3) imply that

$$\lim_{m \rightarrow \infty} \frac{1}{\ln p_m} \sum_{n \leq p_m} \frac{\mu^2(n)}{n} = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} \approx 0.607927. \quad (6)$$

The above result should be compared with the well-known asymptotic density of square-free integers:

$$\lim_{m \rightarrow \infty} \frac{1}{p_m} \sum_{n \leq p_m} \mu^2(n) = \frac{1}{\zeta(2)}. \quad (7)$$

In general, consider an interval $\Delta \subset \mathbb{R} \geq 0$ and denote by $M(\Delta)$ the number of integers $n \in \Delta$ for which $\mu(n) \neq 0$. The usual exclusion–inclusion principle from

Probability Theory gives

$$\left| \frac{M(\Delta)}{|\Delta|} - \frac{1}{\zeta(2)} \right| \leq R(|\Delta|),$$

where $R(L) \rightarrow 0$ as $L \rightarrow \infty$. It is easy to show that $R(L) = O(L^{-1/2})$ and, assuming the Riemann Hypothesis, the best known result is $R(L) = O_\varepsilon(L^{-37/54+\varepsilon})$. See [29] for a survey on this topic.

Comparing (6) and (7), one can see the difference between our ensemble Ω_m , where the probability of n is proportional to $\frac{1}{n}$, and the ensemble $\{n \leq x : \mu(n) \neq 0\}$, where the probability distribution is uniform. Another difference between the two ensembles follows. If we use our Theorem 1 to estimate $M(\Delta)$ when $\Delta = [p_m^{c_1}, p_m^{c_2}]$, $0 < c_1 < c_2$, we get

$$M(\Delta) = \sum_{\substack{n \in \Omega_m \\ c_1 \leq \zeta_m(n) \leq c_2}} 1 = Z_m \sum_{\substack{n \in \Omega_m \\ c_1 \leq \zeta_m(n) \leq c_2}} n \pi_m(n)$$

and therefore

$$\begin{array}{ccc} \frac{e^\gamma \int_{c_1}^{c_2} d(t) dt}{\zeta(2)} \frac{\ln p_m}{p_m^{c_2-c_1}} \sim \frac{Z_m \sum_{c_1 \leq \zeta_m(n) \leq c_2} \pi_m(n)}{(p_m^{c_2-c_1} - 1)} \leq \frac{M(\Delta)}{|\Delta|} \leq \frac{Z_m \sum_{c_1 \leq \zeta_m(n) \leq c_2} \pi_m(n)}{\left(1 - \frac{1}{p_m^{c_2-c_1}}\right)} \sim \frac{e^\gamma \int_{c_1}^{c_2} d(t) dt}{\zeta(2)} \ln p_m \\ \downarrow \qquad \qquad \qquad \downarrow \qquad \qquad \qquad \downarrow \\ 0 \qquad \qquad \qquad \frac{1}{\zeta(2)} \qquad \qquad \qquad \infty \end{array}$$

as $m \rightarrow \infty$. In other words, Theorem 1 can only be used to obtain inequalities concerning $M(\Delta)$, and these inequalities become less precise (and depend on c_1, c_2) as m grows.

Nevertheless, there is a connection between the two ensembles. While Ω_m is very sparse and its largest element is of order m^m , the initial segment $\Omega_m \cap \{1, \dots, p_m\}$ actually coincides with the set $\{n \leq p_m : \mu(n) \neq 0\}$. This fact was used in the analysis of the error term in Theorem 1 for shrinking intervals in [4].

It is worth mentioning that there are several examples in Number Theory where the limiting densities are constant on some intervals. This is the case, for instance, for the gap distribution in the sequence $\{\sqrt{n} \bmod 1\}_{n \in \mathbb{N}}$ studied by Elkies and McMullen [10], and for the distribution of lattice points visible from the origin discovered by Boca et al. [1]. Presumably, this is a general property of a class of infinitely divisible distributions that appear in problems related to Number Theory.

It is well known from Number Theory that the Prime Number Theorem is equivalent to the fact that

$$\sum_{n \leq N} \mu(n) = o(N)$$

as $N \rightarrow \infty$, while the Riemann Hypothesis is equivalent to

$$\sum_{n \leq N} \mu(n) = O_\varepsilon(N^{1/2+\varepsilon})$$

for any $\varepsilon > 0$ as $N \rightarrow \infty$, see [23].

In [31], Sarnak discusses a classical heuristic argument according to which

$$\sum_{n \leq N} \mu(n) \xi(n) = o(N) \quad \text{as } N \rightarrow \infty \quad (8)$$

for any ‘reasonable’ bounded sequence $\xi(n)$, defined independently of $\mu(n)$. If (8) holds, μ is said to be *orthogonal* to ξ . The sequences $\xi(n)$ considered are generated by dynamical systems $S = (X, T)$, where X is a compact topological space, $T : X \rightarrow X$ is continuous, and $\xi(n) = f(T^n x)$ for $x \in X$ and $f \in C(X)$. A system is called *deterministic* if its topological entropy is zero. The sequence $\mu(n)$ is said to be *disjoint* from a system S if $\mu(n)$ is orthogonal to all sequences $\xi(n)$ generated by S as above. Sarnak’s conjecture reads then as: *μ is disjoint from any deterministic system.* Known cases of this conjecture are:

- (i) the trivial case $|X| = 1$ corresponds to the Prime Number Theory;
- (ii) $|X| < \infty$ is covered by Dirichlet’s theorem;
- (iii) if $S = (\mathbb{R}/\mathbb{Z}, R_\alpha)$, where $R_\alpha(x) = x + \alpha$, the conjecture follows from the estimate

$$\max_\alpha \left| \sum_{n \leq N} \mu(n) e^{2\pi i n \alpha} \right| = O_\lambda \left(\frac{N}{(\ln \lambda N)^\lambda} \right)$$

due to Davenport [5];

- (iv) extensions of the latter to Kronecker flows and zero-entropy affine automorphisms (see, e.g. the work by Liu and Sarnak [27]);
- (v) the case of $S = (\Gamma \backslash G, T_\alpha)$, where G is a nilpotent Lie group, Γ a lattice in G , and $T_\alpha(\Gamma x) = \Gamma x \alpha$, addressed by Green and Tao [19].

The conjecture is open for Ratner sequences, i.e. sequences generated as above when G is a semisimple group and α is unipotent. The particular case of $S = (\mathrm{SL}(2, \mathbb{Z}) \backslash \mathrm{SL}(2, \mathbb{R}), T_\alpha)$ is related to the recent work of Sarnak and Ubis [32].

The above conjecture is related to the so-called *Chowla conjecture*, concerning the self-correlation of the Möbius function, that can be phrased as follows: for every $d \in \mathbb{N}$, $0 \leq a_1 \leq a_2 \leq \dots \leq a_d$ and $k_1, k_2, \dots, k_d \in \{1, 2\}$ not all even, we have

$$\sum_{n \leq N} \mu^{k_1}(a_1 + n) \mu^{k_2}(a_2 + n) \dots \mu^{k_d}(a_d + n) = o(N)$$

as $N \rightarrow \infty$. One can show, see [31], that the Chowla conjecture implies the above conjecture. Let us also mention the very recent results by J. Bourgain, showing that Walsh

functions are orthogonal to $\mu(n)$ [2], and by B. Green, who proved that the Möbius sequence is orthogonal to the so-called bounded depth circuit functions, belonging to the class $AC^0(d)$ [18]. Another way to study the randomness $\mu(n)$ is to study the spectral properties of the Schrödinger equations on the positive integers, where $\mu(n)$ plays the role of a “random” potential. It was shown recently by J. Bourgain that almost every eigenfunction of the Schrödinger operator $\Delta + \mu$ on $\ell^2(\mathbb{Z}_{\geq 0})$ has positive Lyapunov exponent [3], in accordance with the theory of (truly) random Schrödinger operators. It is still not known whether Anderson localization holds for the potential $\mu(n)$.

In Sect. 3 we prove the following result. Let $f(t) = \sum_{1 \leq r \leq r_0} c_r t^r$ be an arbitrary polynomial with integer coefficients and $\alpha = \frac{k}{l}$ be a rational number, with $(k, l) = 1$. Let $e(z) = e^{2\pi iz}$

Theorem 2 *For all sufficiently large m*

$$\sum_{n \in \Omega_m} e(\alpha f(n)) \mu(n) = 0. \quad (9)$$

If $f(t) = t$, then Theorem 2 gives an analog of Davenport’s result [5] in the case of our ensemble Ω_m and for rational α . Similarly, if $f(t) = t^2$ we get an analog of a result by Green and Tao for nilflows [19]. The basic difference is that instead of the intervals $[1, N]$ we use the sets Ω_m and (9) is much sharper.

For real α satisfying certain Diophantine conditions some estimates of the sum (9) are given in Sect. 3 for real α satisfying certain Diophantine conditions.

In Sect. 4 we prove a theorem which can be considered as related to the Möbius Randomness Law [23, 31] and sheds some light on the cancellations caused by $(-1)^k$ in the definition of the Möbius function.

Theorem 3 *Let f be a 1-periodic function of the form $f(t) = \sum_{k \in \mathbb{Z}} f_k e(kt)$ and such that $\sum_{k \in \mathbb{Z}} |f_k| < \infty$. Consider*

$$I_m[f] = \sum_{n \in \Omega_m} f\left(\frac{\ln n}{\ln p_m}\right) \mu(n). \quad (10)$$

Then

$$I_m[f] = 0.$$

The proof of Theorem 3 exploits the special structure of our ensemble Ω_m when sampled by functions of the type $f(\frac{\ln \cdot}{\ln p_m})$, where f is 1-periodic. The study of the sum (10) for functions of bigger periods requires the following analysis. For $\frac{k}{r} \in \mathbb{Q}$ let

$$I_m\left(\frac{k}{r}\right) = \sum_{n \in \Omega_m} e\left(\frac{k}{r} \frac{\ln n}{\ln p_m}\right) \mu(n)$$

The following Theorem deals with the case when the period r tends to infinity.

Theorem 4 Let $\frac{k}{r} = \frac{k(m)}{r(m)} \rightarrow 0$ as $m \rightarrow \infty$. Then

$$|I_m(\frac{k}{r})| = \exp\left\{O\left(\frac{m}{\ln m}\right)\right\} \quad (11)$$

as $m \rightarrow \infty$.

Theorem 3 allows to construct a class of functions for which (11) holds. Fix three non-decreasing sequences $\{K(m)\}_{m \in \mathbb{N}}$, $\{r(m)\}_{m \in \mathbb{N}}$, and $\{G(m)\}_{m \in \mathbb{N}}$ such that $K(m), r(m) \rightarrow \infty$, $\frac{K(m)}{r(m)} \rightarrow 0$ as $m \rightarrow \infty$. Consider the space of functions $\mathcal{F}_m = \mathcal{F}_m(K, r, G)$ of all $r(m)$ -periodic functions of the form

$$f^{(m)}(t) = \sum_{k \in \mathbb{Z}} f_k^{(m)} e\left(\frac{k}{r(m)}t\right), \quad f_k^{(m)} \in \mathbb{R},$$

and such that

$$\sum_{|k| \leq K(m)} |f_k^{(m)}| \leq G(m), \quad (12)$$

$$\sum_{|k| > K(m)} |f_k^{(m)}| \leq e^{-m(\ln 2 + O(\frac{1}{\ln m}))} G(m). \quad (13)$$

While the sequence $r(n)$ controls the period of the functions, the estimates involving $K(n)$ and $G(n)$ put some constraint on the decay of the Fourier coefficients $f_k^{(m)}$. We have the following Corollary of Theorem 4:

Corollary 5 Let $f^{(m)} \in \mathcal{F}_m(K, r, G)$ and let

$$I_m[f^{(m)}] = \sum_{n \in \Omega_m} f^{(m)}\left(\frac{\ln n}{\ln p_m}\right) \mu(n).$$

Then

$$|I_m[f^{(m)}]| \leq \exp\left\{O\left(\frac{m}{\ln m}\right)\right\} G(m).$$

In other words, if $G(m)$ does not grow too fast as $m \rightarrow \infty$, Corollary 5 provides a class of functions that are not disjoint (in the sense discussed above) from $\mu(n)$ when sampled along the rescaled ensemble $\{\frac{\ln n}{\ln p_m} : n \in \Omega_m\}$.

In particular, if we want to approximate the indicator function of an interval (whose Fourier coefficient decay as $\frac{1}{k}$), we can choose $K(m) = m^\alpha$, $0 < \alpha < 1$, $r(m) = m$ and $G(m) = c \ln(m)$. In this case the functions

$$f^{(m)}(t) = \sum_{k \in \mathbb{Z}} f_k^{(m)} e\left(\frac{k}{m}t\right) \quad (14)$$

such that

$$\sum_{|k| \leq m^\alpha} |f_k^{(m)}| \leq c \ln m \quad \text{and} \quad \sum_{|k| > m^\alpha} |f_k^{(m)}| \leq 3^{-m}$$

satisfy the estimate

$$|I_m[f^{(m)}]| \leq \exp \left\{ O \left(\frac{m}{\ln m} \right) \right\}. \quad (15)$$

Remark 1 In this case the choice of $r(m) = m$ is motivated by the fact that the largest element of our rescaled ensemble $\left\{ \frac{\ln n}{\ln p_m}, n \in \Omega_m \right\}$ is

$$\frac{\ln(p_1 p_2 \dots p_m)}{\ln p_m} = \frac{(1 + o(1))m \ln m}{\ln p_m} = (1 + o(1))m.$$

The idea is to construct a sequence of functions in $\{f^m\}_{m \in \mathbb{N}}$ that approximate a function supported on $[0, 1]$. Since this interval corresponds to the classical ensemble $\{n \leq p_m, \mu(n) \neq 0\}$, presumably one could be able to get some estimates for the classical problem (8).

There are natural generalizations of the Möbius function. Fix $r \geq 1$ and consider numbers $n = p_1^{v_1} \dots p_m^{v_m}$, $0 \leq v_j \leq r$. The set of such numbers is denoted by $\Omega_m^{(r)}$. It is clear that $|\Omega_m^{(r)}| = (r + 1)^m$ and we can introduce the probability distribution $\Pi_m^{(r)}$ on $\Omega_m^{(r)}$ for which

$$\pi_m^{(r)}(n) = \frac{1}{Z_m^{(r)} \cdot n},$$

where

$$Z_m^{(r)} = \sum_{n \in \Omega_m^{(r)}} \frac{1}{n} \sim \frac{e^\gamma}{\zeta(r + 1)} \ln p_m.$$

The exclusion–inclusion principle (see above) gives the following: if Δ is an arbitrary interval and $M^{(r)}(\Delta) = \#\{n \in \Delta : p_1^{v_1} \dots p_m^{v_m}, 0 \leq v_j \leq r\}$, then

$$\left| \frac{M^{(r)}(\Delta)}{|\Delta|} - \frac{1}{\zeta(r + 1)} \right| \leq R^{(r)}(|\Delta|), \quad (16)$$

where $R^{(r)}(|\Delta|) \rightarrow 0$ as $|\Delta| \rightarrow \infty$. It is easy to see that $R^{(r)}(L) = O(L^{-1 + \frac{1}{r+1}})$. Assuming the Riemann Hypothesis one can get better estimates but they look quite complicated, see again [29] and the references therein.

Theorem 1 is also valid for the ensemble $(\Omega_m^{(r)}, \Pi_m^{(r)})$ with the same limiting probability distribution given by the Dickman–De Bruijn density. Theorem 2 remains true with the same proof.

The sum $\sum_{j=1}^m v_j$ plays an important role in our analysis. For example, $\mu(n) = (-1)^{\sum_{j=1}^m v_j}$. Since terms of the sum are independent—but not identically distributed—random variables, we can use some standard tools from Probability Theory to deduce the asymptotic statistical behavior of the sum. In Sect. 5 we prove the following two limit theorems

Theorem 6 *The sum $\sum_{j=1}^m v_j$ satisfies the following Erdős–Kac Central Limit Theorem: as $m \rightarrow \infty$*

$$\frac{\sum_{j=1}^m v_j - \ln \ln p_m}{\sqrt{\ln \ln p_m}} \implies \mathcal{N}(0, 1)$$

where $\mathcal{N}(0, 1)$ denotes the standard Gaussian distribution with mean 0 and variance 1.

Theorem 7 *For $0 < a < b \leq 1$, the sum*

$$\eta_m^{(a,b)} = \sum_{j=m^a}^{m^b} v_j$$

converges weakly as $m \rightarrow \infty$ to a random variable having Poisson distribution with parameter $\ln \frac{b}{a}$.

Section 5 also discusses some consequences of the theorems above, indicating the asymptotic probability that an element of our ensemble Ω_m has no prime factors larger than certain functions (depending on m), such as

$$\Pi_m \{n = p_1^{v_1} \cdots p_m^{v_m} : v_j = 0 \text{ for } j \geq cm\}, \quad 0 < c \leq 1,$$

or

$$\Pi_m \{n = p_1^{v_1} \cdots p_m^{v_m} : v_j = 0 \text{ for } j \geq m^a\}, \quad 0 < a \leq 1.$$

Section 6 discusses some connections between the Möbius function and the Riemann ζ -function, via the first formula in (1). While in the region $\Re s > 1$ all the sums are absolutely convergent, it is a priori non clear what happens when $0 < \Re s \leq 1$. We introduce a specific way of summing of the sums

$$\sum_{n \geq 1} \frac{\mu(n)}{n^s}, \quad \sum_{n \geq 1} \frac{1}{n^s}, \quad \text{and} \quad \sum_{p \geq 1} \frac{1}{p^s}$$

and study their convergence.

2 The proof of Theorem 1 and some properties of the Dickman–De Bruijn distribution

Proof of Theorem 1 Let $\zeta_m = \sum_{j=1}^m v_j \frac{\ln p_j}{\ln p_m}$. Its characteristic function has the form

$$\begin{aligned}\varphi_m(\lambda) &= \mathbb{E} e^{i\lambda \zeta_m} = \mathbb{E} \exp \left\{ \frac{i\lambda}{\ln p_m} \sum_{j=1}^m v_j \ln p_j \right\} = \prod_{j=1}^m \left(1 + \frac{1}{p_j + 1} \left(e^{\frac{i\lambda \ln p_j}{\ln p_m}} - 1 \right) \right) \\ &= \exp \left\{ \sum_{j=1}^m \ln \left(1 + \frac{1}{p_j + 1} \left(e^{\frac{i\lambda \ln p_j}{\ln p_m}} - 1 \right) \right) \right\}.\end{aligned}$$

Using the function $N(t)$ as before we get

$$\begin{aligned}I_m &= \sum_{j=1}^m \ln \left(1 + \frac{1}{p_j + 1} \left(e^{\frac{i\lambda \ln p_j}{\ln p_m}} - 1 \right) \right) \\ &= \sum_{t=1}^{p_m} (N(t) - N(t-1)) \ln \left(1 + \frac{1}{t+1} \left(e^{\frac{i\lambda \ln t}{\ln p_m}} - 1 \right) \right)\end{aligned}$$

and summation by parts yields

$$I_m = f_m(p_m + 1)N(p_m) + I'_m,$$

where $f_m(s) = \ln \left(1 + \frac{1}{s+1} \left(e^{\frac{i\lambda \ln s}{\ln p_m}} - 1 \right) \right)$ and

$$\begin{aligned}I'_m &= - \sum_{t=1}^{p_m} N(t-1)(f_m(t+1) - f_m(t)) = - \sum_{t=2}^{p_m} \frac{t}{\ln t} \frac{d}{dt} f_m(t) + \text{error terms} \\ &= \sum_{t=2}^{p_m} \frac{t}{\ln t} \frac{\frac{1}{(t+1)^2} \left(e^{\frac{i\lambda \ln t}{\ln p_m}} - 1 \right) - \frac{i\lambda}{t(t+1)\ln p_m} e^{\frac{i\lambda \ln t}{\ln p_m}}}{1 + \frac{1}{t+1} \left(e^{\frac{i\lambda \ln t}{\ln p_m}} - 1 \right)} + \text{error terms} \\ &= \sum_{t=2}^{p_m} \frac{1}{t \ln t} \left(e^{\frac{i\lambda \ln t}{\ln p_m}} - 1 \right) + \text{error terms}.\end{aligned}$$

The analysis of the error terms is left to the interested reader and can be found in [4] and the appendix therein.

Let us consider the change of variables $v = \frac{\ln t}{\ln p_m}$, for which $\frac{dt}{t} = dv \ln p_m$. Then

$$I_m = \sum_v \frac{dv}{v} \left(e^{i\lambda v} - 1 \right) + \text{error terms}$$

and, as $m \rightarrow \infty$, the above sum converges to

$$\int_0^1 \frac{e^{i\lambda v} - 1}{v} dv = \ln \varphi(\lambda).$$

Let us prove that the distribution with characteristic function given by (4) is infinitely divisible. Kolmogorov proved [24] that a probability distribution P_ξ over \mathbb{R} with finite variance is infinitely divisible if and only if its characteristic function $\varphi(\lambda)$ has the form

$$\ln \varphi(\lambda) = i\kappa\lambda + \int_{\mathbb{R}} (e^{i\lambda v} - 1 - i\lambda v) \frac{dK(v)}{v^2}, \quad (17)$$

where κ is a constant and $v \mapsto K(v)$ is a non-decreasing function of bounded variation satisfying $\lim_{v \rightarrow -\infty} K(v) = 0$. It is easy to check that $\kappa = \int_{\mathbb{R}} x dP(x) = \mathbb{E}\xi$ and $\lim_{v \rightarrow \infty} K(v) = \mathbb{E}(\xi - \mathbb{E}\xi)^2$. In our case $\kappa = 1$ and $\lim_{v \rightarrow \infty} K(v) = \frac{1}{2}$ (see Sect. 2.1), and by choosing

$$K(v) = \begin{cases} 0 & \text{if } v < 0; \\ \frac{v^2}{2} & \text{if } 0 \leq v \leq 1; \\ \frac{1}{2} & \text{if } v > 1 \end{cases}$$

in (17) gives (4). Therefore the limiting distribution is infinitely divisible and Theorem 1 is proven. \square

2.1 On the Dickman–De Bruijn distribution

It is known (see [6]) that $\varphi(\lambda)$ is the characteristic function of the *Dickman–De Bruijn distribution*, with density $e^{-\gamma} \rho(t)$, where $\rho(t)$ is determined by the initial condition

$$\rho(t) = \begin{cases} 0, & t \leq 0; \\ 1, & 0 < t \leq 1, \end{cases} \quad (18)$$

and the integral equation

$$t\rho(t) = \int_{t-1}^t \rho(s) ds, \quad t \in \mathbb{R}.$$

The density $e^{-\gamma} \rho(t)$ is plotted in Fig. 1. It also satisfies the delay differential equation

$$t\rho'(t) + \rho(t-1) = 0$$

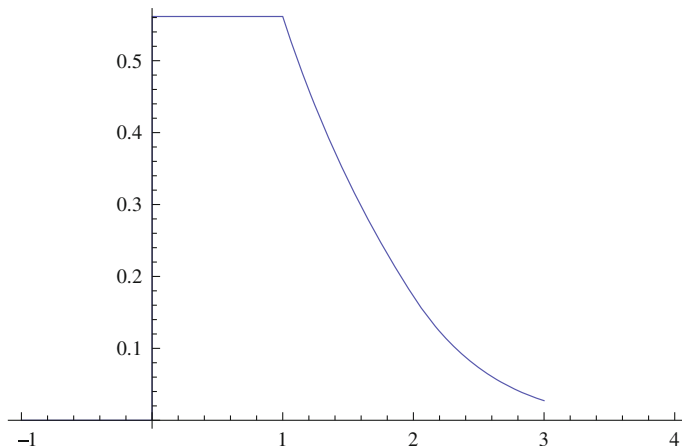


Fig. 1 The density $e^{-\gamma} \rho(t)$ of the Dickman–De Bruijn distribution

for $t \geq 1$ (at $t = 1$ we consider the right derivative) and for every $k = 1, 2, 3, \dots$ there is an analytic function $\rho_k(t)$ that gives $\rho(t)$ on $k - 1 \leq t \leq k$. For example, $\rho_1 \equiv 1$, $\rho_2(t) = 1 - \ln t$ and $\rho_3(t) = 1 - \ln t + \int_2^t \ln(u - 1) \frac{du}{u}$. It is also easy to see that $\rho \in C^k([k, \infty))$ for each k .

Among other properties of $\rho(t)$ one can mention that it is log-concave on $[1, \infty)$ and

$$\rho(t) = \exp \left\{ -t \left(\ln t + \ln \ln t - 1 + \frac{\ln \ln t}{\ln t} + O \left(\frac{(\ln \ln t)^2}{(\ln t)^2} \right) \right) \right\}$$

as $t \rightarrow \infty$. In other words, the limiting density $e^{-\gamma} \rho(t)$ is constant on the interval $(0, 1]$, where it takes the value $e^{-\gamma}$, and then decays faster than exponentially on $(1, \infty)$, like Poisson distribution. In particular, all its moments are finite.

Let us remark that the explicit formula for $\varphi(\lambda)$ allows us to compute the k -th moment m_k of the Dickman–De Bruijn distribution via $m_k = e^{-\gamma} \int_0^\infty t^k \rho(t) dt = \frac{1}{i^k} \frac{d^k}{d\lambda^k} \varphi(\lambda) |_{\lambda=0}$. We get

$$\{m_k\}_{k \geq 1} = \left\{ 1, \frac{3}{2}, \frac{17}{6}, \frac{19}{3}, \frac{81}{5}, \frac{8351}{180}, \frac{184553}{1260}, \frac{52907}{105}, \frac{1768847}{945}, \frac{70442753}{9450}, \frac{1096172081}{34650}, \dots \right\}.$$

The above sequence can therefore be characterized by the following identity:

$$1 + \sum_{k=1}^{\infty} m_k \frac{z^k}{k!} = \exp \left\{ \sum_{n=1}^{\infty} \frac{z^n}{n n!} \right\} = \exp \left\{ \int_0^z \frac{e^t - 1}{t} dt \right\}.$$

Moreover

$$m_k = \sum_{\mathcal{P} \in \mathcal{P}_k} \frac{1}{\prod_{P \in \mathcal{P}} |P|},$$

where \mathcal{P}_k denotes the set of partitions of $\{1, 2, \dots, k\}$, i.e. $\mathcal{P} = \{P_1, P_2, \dots, P_s\} \in \mathcal{P}_k$ satisfies $\bigcup_{j=1}^s P_j = \{1, 2, \dots, k\}$ and $P_i \cap P_j = \emptyset$ for $i \neq j$. For example

$$\mathcal{P}_3 = \{\{\{1, 2, 3\}\}, \{\{1\}, \{2, 3\}\}, \{\{2\}, \{1, 3\}\}, \{\{3\}, \{1, 2\}\}, \{\{1\}, \{2\}, \{3\}\}\}$$

gives $m_3 = \frac{1}{3} + \frac{3}{1 \cdot 2} + \frac{1}{1 \cdot 1 \cdot 1} = \frac{17}{6}$.

Recall the random variable $\zeta_m = \frac{\ln n}{\ln p_m} = \sum_{j=1}^m \nu_j \frac{\ln p_j}{\ln m}$ from Theorem 1. A simple consequence of the fact that $m_1 = 1$ and $m_2 = \frac{3}{2}$ is that $\mathbb{E}\zeta_m \rightarrow 1$ and $\text{Var}\zeta_m = \mathbb{E}\zeta_m^2 - (\mathbb{E}\zeta_m)^2 \rightarrow \frac{3}{2} - 1 = \frac{1}{2}$ as $m \rightarrow \infty$.

The Dickman–De Bruijn density ρ first appeared in the theory of *smooth numbers* (i.e. numbers with small prime factors). Let $\Psi(x, y)$ denote the number of integers $\leq x$ whose prime factors are $\leq y$. Dickman [9] showed that $\Psi(x, x^{1/u}) \sim x\rho(u)$ as $x \rightarrow \infty$. The range of y such that the asymptotic formula $\Psi(x, y) \sim x\rho(u)$, where $x = y^u$, has been significantly enlarged by De Bruijn [6–8] ($y \geq \exp((\ln x)^{5/8+\varepsilon})$) and Hildebrand [22] ($y \geq \exp((\ln \ln x)^{5/3+\varepsilon})$). Notice that in our ensemble Ω_m (where each element is weighted, not simply counted) we have $x = p_1 p_2 \cdots p_m$ and $y = p_m$ and thus $y \sim \ln x$. In this regime Erdős [11] showed that $\ln \Psi(x, \ln x) \sim \frac{\ln 4 \ln x}{\ln \ln x}$ as $x \rightarrow \infty$ and therefore the asymptotic is no longer given by the function ρ . In other words, some kind of *phase transition* occurs in the asymptotic behavior of $\Psi(x, y)$. For a survey on the theoretical and computational aspects of smooth numbers see [16].

Another remarkable occurrence of the Dickman–De Bruijn density comes from the limit of convolution powers, namely

$$e^{-\gamma} \rho(u) = \lim_{n \rightarrow \infty} f_n * f_n * \cdots * f_n(u) \quad (n \text{ times}),$$

where

$$f_n(u) = \begin{cases} \frac{1}{nu} & \text{for } e^{-n} \leq u \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

See [21] for the discussion and generalization of this result.

Goncharov [15] discovered the distribution $e^{-\gamma} \rho(t)$ (although expressed in a somewhat cumbersome form) in 1944 when studying the distribution of maximal cycle length in random permutations. His work, independent of Dickman's, was later popularized by Vershik and Schmidt [35, 36]. The Dickman–De Bruijn distribution also appears when studying the the marginals of the so-called Poisson–Dirichlet distribution, see e.g. [33, 34] and the references therein.

3 Proof of Theorem 2

Proof of Theorem 2 We consider a polynomial $f \in \mathbb{Z}[t]$, $f(t) = \sum_{q=1}^{q_0} c_q t^q$. We want to prove that

$$I_m = \sum_{n \in \Omega_m} e(\alpha f(n)) \mu(n) = 0,$$

for sufficiently large m , when $\alpha = \frac{k}{l} \in \mathbb{Q}$ and $(k, l) = 1$. Let us introduce the sets

$$\mathcal{N}_m = \{p_1, \dots, p_m\}, \quad \mathcal{N}'_m = \{p_j : 1 \leq j \leq m, p_j \equiv 1 \pmod{l}\} \quad \text{and} \\ \mathcal{N}''_m = \mathcal{N}_m \setminus \mathcal{N}'_m.$$

By Dirichlet's theorem, the arithmetic progression $\{1, 1+l, 1+2l, 1+3l, \dots\}$ contains infinitely many primes, and thus $|\mathcal{N}'_m| \rightarrow \infty$ as $m \rightarrow \infty$. Each $n \in \Omega_m$ determines uniquely $\tilde{\mathcal{N}}'_m = \tilde{\mathcal{N}}'_m(n) \subseteq \mathcal{N}'_m$ and $\tilde{\mathcal{N}}''_m = \tilde{\mathcal{N}}''_m(n) \subseteq \mathcal{N}''_m$, where

$$\tilde{\mathcal{N}}'_m = \{p_j \in \mathcal{N}'_m, v_j(n) \neq 0\}, \quad \tilde{\mathcal{N}}''_m = \{p_j \in \mathcal{N}''_m, v_j(n) \neq 0\}, \\ n = \prod_{p' \in \tilde{\mathcal{N}}'_m} p' \cdot \prod_{p'' \in \tilde{\mathcal{N}}''_m} p''.$$

We clearly have

$$\mu(n) = (-1)^{|\tilde{\mathcal{N}}'_m| + |\tilde{\mathcal{N}}''_m|}$$

and thus

$$I_m = \sum_{\tilde{\mathcal{N}}'_m \subseteq \mathcal{N}'_m} \sum_{\tilde{\mathcal{N}}''_m \subseteq \mathcal{N}''_m} e\left(\frac{k}{l} \sum_{1 \leq q \leq q_0} c_q \prod_{p' \in \tilde{\mathcal{N}}'_m} p'^q \cdot \prod_{p'' \in \tilde{\mathcal{N}}''_m} p''^q\right) (-1)^{|\tilde{\mathcal{N}}'_m| + |\tilde{\mathcal{N}}''_m|} \quad (19)$$

Every $p' \in \tilde{\mathcal{N}}'_m$ is of the form $p' = 1 + s'l$ for some integer $s' = s'(p')$, and hence we have $\prod_{p' \in \tilde{\mathcal{N}}'_m} p'^q \equiv 1 \pmod{l}$. This means that

$$\frac{k}{l} \sum_{1 \leq q \leq q_0} c_q \prod_{p' \in \tilde{\mathcal{N}}'_m} p'^q \cdot \prod_{p'' \in \tilde{\mathcal{N}}''_m} p''^q \equiv \frac{k}{l} \sum_{1 \leq q \leq q_0} c_q \prod_{p'' \in \tilde{\mathcal{N}}''_m} p''^q \pmod{1}$$

and does not depend on $\tilde{\mathcal{N}}'_m$. Therefore in (19) we have

$$\begin{aligned} I_m &= \sum_{\tilde{\mathcal{N}}'_m \subseteq \mathcal{N}'_m} \sum_{\tilde{\mathcal{N}}''_m \subseteq \mathcal{N}''_m} e \left(\frac{k}{l} \sum_{1 \leq q \leq q_0} c_q \prod_{p'' \in \tilde{\mathcal{N}}''_m} p''^q \right) (-1)^{|\tilde{\mathcal{N}}'_m|} (-1)^{|\tilde{\mathcal{N}}''_m|} \\ &= \sum_{\tilde{\mathcal{N}}''_m \subseteq \mathcal{N}''_m} e \left(\frac{k}{l} \sum_{1 \leq q \leq q_0} c_q \prod_{p'' \in \tilde{\mathcal{N}}''_m} p''^q \right) (-1)^{|\tilde{\mathcal{N}}''_m|} \cdot \sum_{\tilde{\mathcal{N}}'_m \subseteq \mathcal{N}'_m} (-1)^{|\tilde{\mathcal{N}}'_m|}. \end{aligned}$$

Now, by Dirichlet's theorem, we can find $m^* = m^*(l)$ large enough so that $|\mathcal{N}'_{m^*}| \neq \emptyset$ (and the same of course happens for every $m \geq m^*$). We thus have

$$\sum_{\tilde{\mathcal{N}}'_m \subseteq \mathcal{N}'_m} (-1)^{|\tilde{\mathcal{N}}'_m|} = \sum_{k=0}^{|\mathcal{N}'_m|} \binom{|\mathcal{N}'_m|}{k} (-1)^k = (1 + (-1))^{|\mathcal{N}'_m|} = 0$$

and Theorem 2 is proven. \square

Remark 2 Theorem 2 can be extended to polynomials with rational coefficients $f \in \mathbb{Q}[t]$. Indeed, if $f(t) = \sum_{1 \leq q \leq q_0} c_q t^q$ and $c_q = \frac{a_q}{b_q}$, $(a_q, b_q) = 1$, then we can consider $M = \text{LCM}\{b_q, 1 \leq q \leq q_0\}$ and introduce the polynomial $\tilde{f} = M \cdot f$ and $\tilde{\alpha} = \frac{\alpha}{M}$. In view of this fact, it is important to understand how big m has to be as the denominator of α gets larger. The problem of finding the smallest $m^*(l)$ as in the proof of Theorem 2 can be rephrased as follows: given l , find the least prime $p(l)$ in the arithmetic progression $\{1 + l, 1 + 2l, 1 + 3l, \dots\}$. The problem of estimating how big $p(l)$ can be has a long history and very important contributions were made. In our notation $p(l) = p_{m^*}(l)$.

As far as a lower bound is concerned, the Prime Number Theorem implies that $p(l) \geq (1 + o(1))\phi(l) \ln l$ as $l \rightarrow \infty$, where ϕ denotes the Euler's totient function. The factor $(1 + o(1))$ has been improved to $(e^\gamma + o(1))$ by Pomerance [30]. It is also known that if l has at most $\exp(\frac{\ln \ln l}{\ln \ln \ln l})$ prime factors (as happens for 'almost all' l), then one has the bound $p(l) \geq (e^\gamma + o(1))\phi(l) \frac{(\ln \ln l)(\ln \ln \ln l)}{(\ln \ln \ln l)^2}$. It was conjectured by Granville and Pomerance [17] that $p(l) \gg \phi(l) \ln^2 l$.

Concerning the upper bounds, Linnik [25, 26] proved that

$$p(l) \ll l^L. \quad (20)$$

The constant L (called *Linnik's constant*) has been estimated by several authors by exploiting the connection with zero-free regions for Dirichlet L-functions. The current record is due to Xylouris [37], which gives $L \leq 5.2$. It is conjectured by Heath-Brown [20] that $p(l) \ll l^2$. Assuming the Generalized Riemann Hypothesis, one can show that $p(l) \ll \phi(l)^2 \ln^2 l$.

Remark 3 Theorem 2 can be extended to more general ensembles as follows. For $m_1 \leq m_2$ let Ω_{m_1, m_2} be the set of numbers n of the form $n = p_{m_1}^{v_{m_1}} \cdots p_{m_2}^{v_{m_2}}$, where $v_j \in \{0, 1\}$. In particular $\Omega_m = \Omega_{1, m}$. In this case

$$I_{m_1, m_2} = \sum_{n \in \Omega_{m_1, m_2}} e(\alpha f(n)) \mu(n) = 0 \quad (21)$$

whenever there exists a prime $p = p(l, m_1, m_2)$ such that $p_{m_1} \leq p \leq p_{m_2}$ and $p \equiv 1 \pmod{l}$.

Let us restrict our analysis to the case $f(t) = t^d$. In order to study the sum in (9) for $\alpha \in \mathbb{R}$, we need to assume certain Diophantine properties. We have the following

Proposition 8 Let $\alpha \in \mathbb{R}, m \in \mathbb{N}, \frac{k}{l} \in \mathbb{Q}$ satisfy the inequality

$$\left| \alpha - \frac{k}{l} \right| \leq \frac{1}{\tau \cdot (p_1 \cdots p_m)^d} \quad (22)$$

for some $\tau \geq 1$, and assume that

$$\text{there exists } j \leq m \text{ such that } p_j \equiv 1 \pmod{l}. \quad (23)$$

Then

$$\left| \sum_{n \in \Omega_m} e(\alpha n^d) \mu(n) \right| \leq \frac{2^m}{\tau/2\pi} \left(1 + \frac{\pi}{\tau} \right).$$

Notice that the above inequality improves the trivial bound $|\sum_{m \in \Omega_m} e(\alpha n^d) \mu(n)| \leq 2^m$ as long as $\tau > (1 + \sqrt{3})\pi \approx 8.583$.

Proof of Proposition 8 We write

$$\sum_{n \in \Omega_m} e(\alpha n^d) \mu(n) = \sum_{n \in \Omega_m} e\left(\left(\alpha - \frac{k}{l}\right) n^d\right) e\left(\frac{k}{l} n^d\right) \mu(n)$$

and

$$e\left(\left(\alpha - \frac{k}{l}\right) n^d\right) = 1 + 2\pi i \left(\alpha - \frac{k}{l}\right) n^d + R(n),$$

where, by assumption (22),

$$\left| \alpha - \frac{k}{l} \right| n^d \leq \frac{1}{\tau} \quad \text{and} \quad |R(n)| \leq \frac{2\pi^2}{\tau^2}.$$

Thus, we get

$$\left| \sum_{n \in \Omega_m} e(\alpha n^d) \mu(n) \right| \leq \left| \sum_{n \in \Omega_m} e\left(\frac{k}{l} n^d\right) \mu(n) \right| + 2\pi \left| \alpha - \frac{k}{l} \right| \left| \sum_{n \in \Omega_m} n^d e\left(\frac{k}{l} n^d\right) \mu(n) \right| + \left| \sum_{n \in \Omega_m} R(n) e\left(\frac{k}{l} n^d\right) \mu(n) \right| \quad (24)$$

Observe that first sum in (24) is zero by our second hypothesis. The second sum in (24) is estimated as follows

$$\left| \alpha - \frac{k}{l} \right| \left| \sum_{n \in \Omega_m} n^d e\left(\frac{k}{l} n^d\right) \mu(n) \right| \leq \frac{1}{\tau \cdot (p_1 \cdots p_m)^d} \sum_{n \in \Omega_m} n^d \leq \frac{2^m}{\tau}, \quad (25)$$

while the third sum in (24) gives simply

$$\left| \sum_{n \in \Omega_m} R(n) e\left(\frac{k}{l} n^d\right) \mu(n) \right| \leq \sum_{n \in \Omega_m} |R(n)| \leq \frac{2^m}{\tau^2 / 2\pi^2}. \quad (26)$$

Combining (25) and (26) we get the desired estimate. \square

We can now use Proposition 8 in order to obtain an analog of (8) for the ensemble Ω_m for $\xi(n) = e(\alpha n^d)$ when $\alpha \in \mathbb{R}$. We have the following

Corollary 9 *Let α satisfy the following assumptions: there exists a sequence $\{k_m/l_m\}_m \subseteq \mathbb{Q}$ such that*

(i) *There exists $\varepsilon > 0$ such that for every sufficiently large m we have*

$$\frac{\ln \left| \alpha - \frac{k_m}{l_m} \right|}{-m \ln m} \geq d + \varepsilon; \quad (27)$$

(ii) *for sufficiently large m the denominators $\{l_m\}_m$ satisfy*

$$l_m \ll (m \ln m)^{1/L}, \quad (28)$$

where L is Linnik's constant (see Remark 2) and the implied constant is the same as in (20).

Then

$$\left| \sum_{n \in \Omega_m} e(\alpha n^d) \mu(n) \right| = o(2^m). \quad (29)$$

The assumptions of the above Corollary put a rather stringent condition on the type of $\alpha \in \mathbb{R}$ that can be considered. The requirement is that α can be super-exponentially well approximated by a sequence of rational numbers whose denominators do not grow too fast. The explicit dependence on Linnik's constant L is remarkable. In fact, any improvement of the known estimate ($L \leq 5.2$) by Xylouris [37] towards the value $L = 2$ conjectured by Heath-Brown [20] would enlarge the set of α affected by the Corollary.

Proof of Corollary 9 The assumption (27) implies that for every $\varepsilon' \leq \varepsilon$ we have

$$\left| \alpha - \frac{k_m}{l_m} \right| \leq \frac{1}{e^{(d+\varepsilon')m \ln m}} \leq \frac{1}{e^{(d+o(1))m \ln m}},$$

where the second inequality holds for sufficiently large m for every function $o(1)$ as $m \rightarrow \infty$

Let us consider an arbitrary sequence $\{\tau_m\}_m$ such that $\tau_m \rightarrow \infty$ and $\ln \tau_m = o(m \ln m)$ as $m \rightarrow \infty$. Then the sequence $(\tau_m \cdot (p_1 \cdots p_m)^d)^{-1}$ satisfies the inequality

$$\frac{1}{\tau_m \cdot (p_1 \cdots p_m)^d} = \frac{1}{e^{(d+o(1)+\frac{\ln \tau_m}{m \ln m})m \ln m}} \geq \left| \alpha - \frac{k_m}{l_m} \right|$$

for sufficiently large m . Therefore we have the condition (22).

On the other hand, the assumption (28) implies (23). In fact, the asymptotic inequality $l_m \lesssim (\frac{1}{C} m \ln m)^{1/L}$ as $m \rightarrow \infty$ is equivalent to $p_m \gtrsim C l_m^L$, and therefore $p_m \geq p(l_m) = p_m^*(l_m)$ (see the notation used above) and $p(l_m) \equiv 1 \pmod{l_m}$. Now, by Proposition 8 we have

$$\left| \sum_{n \in \Omega_m} e(\alpha n^d) \mu(n) \right| \leq \frac{2^m}{\frac{1}{2\pi} \tau_m} \left(1 + \frac{\pi}{\tau_m} \right) = o(2^m)$$

as $m \rightarrow \infty$, which is precisely what we aimed to prove. \square

4 Proofs of Theorems 3 and 4

Proof of Theorem 3 Since $f(t) = \sum_{k \in \mathbb{Z}} f_k e(kt)$, it is enough to show that

$$I_m(k) = \sum_{n \in \Omega_m} e\left(k \frac{\ln n}{\ln p_m}\right) \mu(n) = 0.$$

Since $\ln n = \sum_{j=1}^m v_j \ln p_j$ and $\mu(n) = (-1)^{\sum_{j=1}^m v_j}$, we can write

$$\begin{aligned} I_m(k) &= \sum_{v_1, \dots, v_m} e\left(k \sum_{j=1}^m v_j \frac{\ln p_j}{\ln p_m}\right) (-1)^{\sum_{j=1}^m v_j} = \prod_{j=1}^m \sum_{v_j \in \{0,1\}} e\left(k \frac{\ln p_j}{\ln p_m} v_j\right) (-1)^{v_j} \\ &= \prod_{j=1}^m \left(1 - e\left(k \frac{\ln p_j}{\ln p_m}\right)\right) = 0 \end{aligned} \quad (30)$$

since for $j = m$ we get a zero factor in the product in (30). \square

Another way to see that $I_m(k) = 0$ is the following:

$$\begin{aligned} I_m(k) &= \sum_{v_1, \dots, v_{m-1}} \left(e\left(k \sum_{j=1}^{m-1} v_j \frac{\ln p_j}{\ln p_m}\right) (-1)^{\sum_{j=1}^{m-1} v_j} \right. \\ &\quad \left. + e\left(k \sum_{j=1}^{m-1} v_j \frac{\ln p_j}{\ln p_m} + k\right) (-1)^{1+\sum_{j=1}^{m-1} v_j} \right) = 0. \end{aligned}$$

In other words, the ensemble Ω_m can be seen as two copies of the ensemble Ω_{m-1} (corresponding to the values of $v_m = \pm 1$), and these two contribute to the sum $I_m(k)$ in opposite ways.

Proof of Theorem 4 We use again the special structure of our ensemble Ω_m for which the sum $I_m(\frac{k}{r})$ can be written as a product.

$$I_m\left(\frac{k}{r}\right) = \sum_{n \in \Omega_m} e\left(\frac{k}{r} \frac{\ln n}{\ln p_m}\right) \mu(n) = \prod_{j=1}^m \left(1 - e\left(\frac{k}{r} \frac{\ln p_j}{\ln p_m}\right)\right) = \exp\{J_m(\frac{k}{r})\},$$

where

$$\begin{aligned} J_m\left(\frac{k}{r}\right) &= \sum_{j=1}^m \ln\left(1 - e\left(\frac{k}{r} \frac{\ln p_j}{\ln p_m}\right)\right) = \sum_{t=1}^{p_m} (N(t) - N(t-1)) \ln\left(1 - e\left(\frac{k}{r} \frac{\ln t}{\ln p_m}\right)\right) \\ &= N(p_m) \ln\left(1 - e\left(\frac{k}{r} \frac{\ln(p_m+1)}{\ln p_m}\right)\right) + J'_m\left(\frac{k}{r}\right) \end{aligned} \quad (31)$$

and, by summation by parts,

$$J'_m\left(\frac{k}{r}\right) = - \sum_{t=1}^{p_m} N(t-1) \left(\ln\left(1 - e\left(\frac{k}{r} \frac{\ln(t+1)}{\ln p_m}\right)\right) - \ln\left(1 - e\left(\frac{k}{r} \frac{\ln t}{\ln p_m}\right)\right) \right). \quad (32)$$

Notice that the boundary term in (31) satisfies the asymptotic estimate

$$\begin{aligned}
 N(p_m) \ln \left(1 - e \left(\frac{k}{r} \frac{\ln(p_m + 1)}{\ln p_m} \right) \right) &= m \ln \left(1 - e \left(\frac{k}{r} \right) - \frac{2\pi i \frac{k}{r} e \left(\frac{k}{r} \right)}{p_m \ln p_m} + O \left(\frac{1}{p_m^2} \right) \right) \\
 &= m \ln \left(1 - e \left(\frac{k}{r} \right) \right) - \frac{2\pi i \frac{k}{r} e \left(\frac{k}{r} \right)}{1 - e \left(\frac{k}{r} \right)} \frac{m}{p_m \ln p_m} + O \left(\frac{m}{p_m^2} \right) \\
 &= m \ln \left(1 - e \left(\frac{k}{r} \right) \right) + O \left(\frac{1}{\ln^2 m} \right) \\
 &= m \left(\ln \left(\frac{-2\pi i k}{r} \right) + \frac{\pi i k}{r} + O \left(\frac{1}{r^2} \right) \right) + O \left(\frac{1}{\ln^2 m} \right)
 \end{aligned} \tag{33}$$

Since the imaginary part of $J_m(\frac{k}{r})$ does not play any role in the estimate of $I_m(\frac{k}{r})$, we have from (33) that

$$\Re \left(m \ln \left(1 - e \left(\frac{k}{r} \frac{\ln(p_m + 1)}{\ln p_m} \right) \right) \right) = m (\ln(2\pi |k|) - \ln r) + O \left(\frac{m}{r^2} \right) + O \left(\frac{1}{\ln^2 m} \right)$$

and therefore the boundary term tends to $-\infty$ (and does not contribute in the estimate for $I_m(\frac{k}{r})$) as long as $\frac{k}{r} \rightarrow 0$ as $m \rightarrow \infty$. The sum (32) is estimated as follows

$$\begin{aligned}
 J'_m \left(\frac{k}{r} \right) &= - \sum_{t=1}^{p_m} N(t-1) \frac{d}{dt} \ln \left(1 - e \left(\frac{k}{r} \frac{\ln t}{\ln p_m} \right) \right) + \text{error terms} \\
 &= - \sum_{t=2}^{p_m} \frac{t}{\ln t} \frac{-e \left(\frac{k}{r} \frac{\ln t}{\ln p_m} \right) \frac{2\pi i \frac{k}{r} \frac{1}{t}}{\ln p_m}}{1 - e \left(\frac{k}{r} \frac{\ln t}{\ln p_m} \right)} + \text{error terms} \\
 &= \frac{2\pi i \frac{k}{r}}{\ln p_m} \sum_{t=2}^{p_m} \frac{1}{\ln t} \frac{e \left(\frac{k}{r} \frac{\ln t}{\ln p_m} \right)}{1 - e \left(\frac{k}{r} \frac{\ln t}{\ln p_m} \right)} + \text{error terms} \\
 &= O(1) \sum_{t=2}^{p_m} \frac{1}{\ln^2 t} + \text{error terms} = O \left(\frac{p_m}{\ln^2 p_m} \right) = O \left(\frac{m}{\ln m} \right).
 \end{aligned}$$

This concludes the proof of Theorem 4. \square

Remark 4 Estimates of the higher order terms in the proof of Theorem 4 could provide a more precise result. Moreover, with a more careful analysis of cancellations, one can presumably handle the case when $\frac{k}{r}$ is simply bounded.

The proof of Corollary 5 is simple and is a direct application of Theorem 4.

Proof of Corollary 5 Let us write

$$I_m[f^{(m)}] = I_m[f^{(m)}] + I_m^{(2)}[f^{(m)}],$$

where

$$I_m[f^{(m)}] = \sum_{|k| \leq K(m)} f_k^{(m)} \sum_{n \in \Omega_m} e\left(\frac{k}{r(m)} \frac{\ln n}{\ln p_m}\right) \mu(n),$$

$$I_m^{(2)}[f^{(m)}] = \sum_{|k| > K(m)} f_k^{(m)} \sum_{n \in \Omega_m} e\left(\frac{k}{r(m)} \frac{\ln n}{\ln p_m}\right) \mu(n).$$

The first sum is estimated using Theorem 4 and (12):

$$\left| I_m(f^{(m)}) \right| \leq \sum_{|k| \leq K(m)} |f_k^{(m)}| \cdot \left| I_m\left(\frac{k}{r(m)}\right) \right| \leq \exp\left\{O\left(\frac{m}{\ln m}\right)\right\} G(m),$$

while the second sum is estimated using (13):

$$|I_m^{(2)}(f^{(m)})| \leq \sum_{|k| > K(m)} |f_k^{(m)}| 2^m \leq \exp\left\{O\left(\frac{m}{\ln m}\right)\right\} G(m).$$

□

5 Distribution of $\sum v_j$

Let us compute the expectation and the variance (with respect to Π_m) of the sum of the independent (but not identically distributed) random variables v_j :

$$\begin{aligned} \mathbb{E}\left(\sum_{j=1}^m v_j\right) &= \sum_{j=1}^m \mathbb{E}v_j = \sum_{j=1}^m \frac{1}{1+p_j} \sim \ln \ln p_m; \\ D_m^2 &= \text{Var}\left(\sum_{j=1}^m v_j\right) = \sum_{j=1}^m \text{Var}v_j = \sum_{j=1}^m \left(\frac{1}{1+p_j} - \frac{1}{(1+p_j)^2}\right) \\ &= \sum_{j=1}^m \frac{p_j}{(1+p_j)^2} \sim \ln \ln p_m. \end{aligned}$$

Proof of Theorem 6 It is easy to see that the sum $\sum_{j=1}^m v_j$ satisfies the Lindeberg condition

$$\lim_{m \rightarrow \infty} \frac{1}{D_m^2} \sum_{j=1}^m \int_{\{x: |x - \mathbb{E}v_j| \geq \varepsilon D_m\}} (x - \mathbb{E}v_j)^2 dF_{v_j}(x) = 0. \quad (34)$$

In fact, since $dF_{v_j}(x) = \left(\frac{p_j}{1+p_j} \delta_0 + \frac{1}{1+p_j} \delta_1\right) dx$ and $D_m \rightarrow \infty$ as $m \rightarrow \infty$, then for sufficiently large m we have that $x \leq \frac{1}{1+p_j} - \varepsilon D_m < 0$ and $x \geq \frac{1}{1+p_j} + \varepsilon D_m > 1$,

and thus the m integrals in (34) are identically zero. The Lindeberg condition implies the Central Limit Theorem for $\sum_{j=1}^m v_j$. Notice that this result is an analog of the celebrated Erdős–Kac Central Limit Theorem [12]. \square

Proof of Theorem 7 Let us compute the characteristic function of $\eta_m^{(a,b)}$ and then take the limit as $m \rightarrow \infty$:

$$\begin{aligned}\varphi_{\eta_m^{(a,b)}}(\lambda) &= \mathbb{E} e^{i\lambda \sum_{j=m^a}^{m^b} v_j} = \prod_{j=m^a}^{m^b} \mathbb{E} e^{i\lambda v_j} = \prod_{j=m^a}^{m^b} \left(\frac{p_j + e^{i\lambda}}{1 + p_j} \right) \\ &= \exp \left\{ \sum_{j=m^a}^{m^b} \ln \left(\frac{p_j + e^{i\lambda}}{1 + p_j} \right) \right\} \\ &= \exp \left\{ \sum_{j=m^a}^{m^b} \ln \left(1 + \frac{e^{i\lambda} - 1}{1 + p_j} \right) \right\} = \exp \left\{ \sum_{j=m^a}^{m^b} \frac{e^{i\lambda} - 1}{1 + p_j} + o(1) \right\} \\ &= \exp \left\{ (e^{i\lambda} - 1) (\ln \ln p_{m^b} - \ln \ln p_{m^a-1}) + o(1) \right\} \\ &= \exp \left\{ (e^{i\lambda} - 1) \ln \left(\frac{\ln p_{m^b}}{\ln p_{m^a}} \right) + o(1) \right\} \rightarrow \exp \left\{ (e^{i\lambda} - 1) \ln \frac{b}{a} \right\}\end{aligned}$$

as $m \rightarrow \infty$. Since the characteristic function of $\text{Pois}(\sigma)$ is $e^{\sigma(e^{i\lambda}-1)}$, we obtain, as $m \rightarrow \infty$,

$$\eta_m^{(a,b)} \Rightarrow \text{Pois}(\ln \frac{b}{a}).$$

\square

Notice that if we replace m^a (or m^b) by $c_1 m^a$ (or $c_2 m^b$), we do not affect the limit distribution. In particular, when $a = b = 1$ we get a degenerate distribution concentrated at 0:

$$\Pi_m \left\{ \sum_{j=c_1 m}^{c_2 m} v_j = 0 \right\} \rightarrow 1 \quad (35)$$

for every $c_1 < c_2$ as $m \rightarrow \infty$. The exponential rate of convergence in (35) will depend on c_2/c_1 and can be easily obtained:

$$\begin{aligned}\ln \left(\Pi_m \left\{ \sum_{j=c_1 m}^{c_2 m} v_j = 0 \right\} \right) &= \ln \left(\prod_{j=c_1 m}^{c_2 m} \frac{p_j}{1 + p_j} \right) = \sum_{j=c_1 m}^{c_2 m} \ln \left(1 - \frac{1}{1 + p_j} \right) \\ &= - \sum_{j=c_1 m}^{c_2 m} \frac{1}{p_j} + o(1)\end{aligned}$$

$$\begin{aligned}
&= -\ln\left(\frac{\ln p_{c_2 m}}{\ln p_{c_1 m}}\right) + o(1) \sim -\ln\left(1 - \frac{\ln(c_2/c_1)}{\ln c_1 + \ln m}\right) \\
&\sim \frac{\ln(c_2/c_1)}{\ln m}
\end{aligned}$$

as $m \rightarrow \infty$. This implies (for $0 < c_1 < c_2 \leq 1$) that

$$\begin{aligned}
&\ln m \cdot \ln(\Pi_m\{n = p_1^{v_1} \cdots p_m^{v_m} : v_j = 0 \text{ for } c_1 m \leq j \leq c_2 m\}) \\
&= \ln m \cdot \ln\left(\Pi_m\left\{\sum_{j=c_1 m}^{c_2 m} v_j = 0\right\}\right) \rightarrow \ln\left(\frac{c_2}{c_1}\right)
\end{aligned}$$

as $m \rightarrow \infty$. In other words (taking $c_1 = \frac{1}{2}$ and $c_2 = 1$), for large m , the probability that n has no prime factors larger than $p_{m/2} \sim \frac{m \ln m}{2}$ is approximately $2^{1/\ln m}$, i.e. exponentially close to 1. In other words:

$$\log(\Pi_m(\Omega_{\lfloor m/2 \rfloor})) \sim \frac{\ln 2}{\ln m} \quad (36)$$

as $m \rightarrow \infty$. On the other hand, a straightforward application of Theorem 7 is gives the following asymptotic probability: for $0 < a < b \leq 1$

$$\begin{aligned}
&\Pi_m\{n = p_1^{v_1} \cdots p_m^{v_m} : v_j = 0 \text{ for } m^a \leq j \leq m^b\} \\
&= \Pi_m\left\{\sum_{j=m^a}^{m^b} v_j = 0\right\} \rightarrow e^{-\ln\left(\frac{b}{a}\right)} = \frac{a}{b}
\end{aligned}$$

as $m \rightarrow \infty$. For example, for $a = \frac{1}{2}$ and $b = 1$, we obtain the remarkable fact that the probability that n has no prime factors larger than $p_{\sqrt{m}} \sim \frac{\sqrt{m} \ln m}{2}$ tends to $\frac{1}{2}$ as $m \rightarrow \infty$. Equivalently,

$$\Pi_m(\Omega_{\lfloor \sqrt{m} \rfloor}) \sim \frac{1}{2} \quad (37)$$

as $m \rightarrow \infty$. The asymptotic relations (36) and (37) show precisely how the measure Π_m is concentrated on square-free numbers with small prime factors.

Large deviations

The probability of large deviations for the sum $\sum_j v_j$ can be also studied. K. Mody proved the following result (K. Mody, private communication):

Proposition 10 *For every $c > 1$ and every $\varepsilon, \varepsilon' > 0$*

$$\Pi_m \left\{ \sum_{j=1}^m v_j \geq c \ln \ln p_m \right\} \leq \left(\frac{\zeta(2)}{e^{\varepsilon'}} + \varepsilon \right) \frac{1}{t^{1-\varepsilon'} \ln t},$$

where $t = k^k$ and $k = \lceil c \log \log p_m \rceil$.

One can check that in Proposition 10 we have, as $m \rightarrow \infty$,

$$\ln(t) = \ln(k^k) \sim c(\ln \ln m)(\ln \ln \ln m)$$

and this provides a sub-exponential bound for the probability of large deviations.

6 Connection with the Riemann ζ -function

In this section we investigate the connection between our ensemble Ω_m and the identity

$$\sum_{n \geq 1} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}. \quad (38)$$

For $\Re s > 1$ both sides of (38) are absolutely convergent series and one can compute the series on the left-hand-side as

$$\sum_{n \geq 1} \frac{\mu(n)}{n^s} = \lim_{m \rightarrow \infty} \sum_{n \in \Omega_m} \frac{\mu(n)}{n^s}. \quad (39)$$

However, for $0 < \Re s \leq 1$, if the series in the LHS of (39) converges, then it is conditionally convergent and, a priori, the limits

$$\kappa(s) = \lim_{m \rightarrow \infty} \sum_{n \in \Omega_m} \frac{\mu(n)}{n^s} \quad \text{and} \quad \lim_{N \rightarrow \infty} \sum_{n=1}^N \frac{\mu(n)}{n^s} \quad (40)$$

can be different.

Let us write $s = s_1 + i s_2$. First, we discuss the case when $s = 1 + i s_2$ and, for simplicity $s_2 > 0$. We prove the following

Lemma 11 *For every $s_2 > 0$ there exists $\tau = \tau(s_2) > 1$ such that, for some constant $C > 0$ and all N ,*

$$\left| \sum_{n=N}^{\lfloor \tau N \rfloor} \frac{1}{n^{1+i s_2}} \right| \leq \frac{C}{N}. \quad (41)$$

Proof Introduce the variable $t = \frac{n}{N}$ for which $dt = \frac{1}{N}$. Then

$$\begin{aligned} \sum_{n=N}^{\lfloor \tau N \rfloor} \frac{1}{n^{1+is_2}} &= \frac{1}{N^{is_2}} \sum_{t=1}^{\tau} \frac{dt}{t^{1+is_2}} + O\left(\frac{1}{N}\right) \sim \frac{1}{N^{is_2}} \int_1^{\tau} \frac{dt}{t^{1+is_2}} + O\left(\frac{1}{N}\right) \\ &= \frac{i(e^{-is_2 \ln \tau} - 1)}{s_2 N^{is_2}} + O\left(\frac{1}{N}\right) \end{aligned} \quad (42)$$

If we choose $\tau = \tau(s_2) = e^{2\pi/s_2} > 1$, the contribution from the integral (42) is zero and we obtain the desired estimate. \square

Notice that if $s_2 < 0$, we can choose $\tau = \tau(s_2) = e^{-2\pi/s_2} > 1$ and the above argument still works.

The above Lemma gives some way to sum the (conditionally convergent) series $\sum_{n \geq 1} \frac{1}{n^s}$. Namely, let $\tau = \tau(s_2)$ be as above and fix $N_0 \in \mathbb{N}$. Define $N_1 = \lfloor \tau N_0 \rfloor + 1$, $N_2 = \lfloor \tau N_1 \rfloor + 1, \dots, N_k = \lfloor \tau N_{k-1} \rfloor + 1$. We have the following

Proposition 12 *Let $s = 1 + is_2$, $s_2 \neq 0$, and let the sequence $\{N_k\}_k$ be defined as above. Then the limit*

$$\gamma(s) = \lim_{k \rightarrow \infty} \sum_{n=1}^{N_k} \frac{1}{n^{1+s_2}} \quad (43)$$

exists.

Proof The proof follows immediately from (41) after writing

$$\sum_{n=N_0}^{N_k-1} \frac{1}{n^{1+is_2}} = \sum_{l=0}^{k-1} \sum_{n=N_l}^{\lfloor \tau N_l \rfloor} \frac{1}{n^{1+is_2}}$$

and observing that, since $\tau(s_2) > 1$, the sequence $\{N_k\}_k$ grows exponentially fast as $k \rightarrow \infty$. \square

The case $s = 1$ (not covered by the previous result) can be treated in more generality. We have the following result for real s in the real segment $(0, 1]$:

Proposition 13 *For $0 < s \leq 1$*

$$\lim_{m \rightarrow \infty} \sum_{n \in \Omega_m} \frac{\mu(n)}{n^s} = 0 \quad (44)$$

Proof For $s = 1$ we have the estimate

$$\begin{aligned} \sum_{n \in \Omega_m} \frac{\mu(n)}{n} &= Z_m \sum_{n \in \Omega_m} \mu(n) \pi_m = Z_m \mathbb{E}(-1)^{\sum_{j=1}^m v_j} = Z_m \prod_{j=1}^m \mathbb{E}(-1)^{v_j} \\ &= Z_m \prod_{j=1}^m \left(\frac{p_j}{1+p_j} - \frac{1}{1+p_j} \right) = Z_m \exp \left\{ \sum_{j=1}^m \ln \left(1 - \frac{2}{1+p_j} \right) \right\} \\ &= Z_m \exp \left\{ -2 \sum_{j=1}^m \frac{1}{p_j} + O(1) \right\} = Z_m \exp \{ -2 \ln \ln p_m + O(1) \} \\ &= \frac{O(1) Z_m}{\ln^2 p_m} \sim \frac{O(1)}{\ln p_m} \end{aligned}$$

and therefore

$$\lim_{m \rightarrow \infty} \sum_{n \in \Omega_m} \frac{\mu(n)}{n} = 0.$$

For $0 < s < 1$ the sum in (44) cannot be interpreted as an expectation with respect to the probability measure Π_m . However, let $\frac{1}{k+1} < s \leq \frac{1}{k}$ for some $k \in \mathbb{N}$. Then

$$\sum_{n \in \Omega_m} \frac{\mu(n)}{n^s} = \prod_{j=1}^m \left(1 - \frac{1}{p_j^s} \right) = \exp \left\{ \sum_{j=1}^m \ln \left(1 - \frac{1}{p_j^s} \right) \right\}.$$

The first k terms of the Taylor expansion of $\ln \left(1 - \frac{1}{p_j^s} \right)$ give divergent series, but these series are dominated by the first term:

$$\begin{aligned} \sum_{n \in \Omega_m} \frac{\mu(n)}{n^s} &= \exp \left\{ - \sum_{j=1}^m \frac{1}{p_j^s} - \frac{1}{2} \sum_{j=1}^m \frac{1}{p_j^{2s}} - \cdots - \frac{1}{k} \sum_{j=1}^m \frac{1}{p_j^{ks}} + O(1) \right\} \\ &\leq \exp \left\{ -C \sum_{j=1}^m \frac{1}{p_j^s} \right\} \leq \exp \{ -C \ln \ln p_m + O(1) \} = \frac{O(1)}{\ln^C p_m} \end{aligned}$$

for sufficiently large m , where $C > 0$. The result is thus proven. \square

Let us remark that (44) is consistent with the fact that $\zeta(s)$ has a simple pole at $s = 1$. For $0 < s < 1$, the function $\kappa(s)$ does not agree with (the analytical continuation of) $\frac{1}{\zeta(s)}$ but it is consistent with the fact that the series $\sum_{n \geq 1} \frac{1}{n^s}$ diverges.

Let us return to the analysis of the sum in (44) when s is on the boundary of absolute convergence of (38), i.e. $s = 1 + is_2$. The following Proposition shows the existence of the limit of such sums and relates them to the so-called Prime Zeta Function $\zeta_P(s) = \sum_{p \geq 1} \frac{1}{p^s}$.

Proposition 14 *The limit*

$$\kappa(s) = \lim_{m \rightarrow \infty} \sum_{n \in \Omega_m} \frac{\mu(n)}{n^{1+is_2}}$$

exists for all $s_2 \neq 0$.

Proof Let us assume $s > 0$. We have the following

$$\sum_{n \in \Omega_m} \frac{\mu(n)}{n^{1+is_2}} = \prod_{j=1}^m \left(1 - \frac{1}{p_j^{1+is_2}}\right) = \exp \left\{ - \sum_{j=1}^m \frac{1}{p_j^{1+is_2}} + O(1) \right\} \quad (45)$$

For the truncation of the Prime Zeta Function we can write

$$\begin{aligned} \sum_{j=1}^m \frac{1}{p_j^{1+is_2}} &= \sum_{t=1}^{p_m} (N(t) - N(t-1)) \frac{1}{t^{1+is_2}} = \sum_{t=1}^{p_m} N(t) \frac{1+is_2}{t^{2+is_2}} + \text{error terms} \\ &= \sum_{t=1}^{p_m} \frac{1}{\ln t} \frac{1+is_2}{t^{1+is_2}} + \text{error terms}. \end{aligned}$$

Consider $M_0 \in \mathbb{N}$ and define $M_k = 2^k M_0$, $k = 1, 2, \dots$. Setting $t = \tau M_k$ we have $d\tau = \frac{1}{M_k}$

$$\begin{aligned} \sum_{M_k \leq t < M_{k+1}} \frac{1}{\ln t} \frac{1}{t^{1+is_2}} &= \frac{1}{M_k^{is_2}} \sum_{1 \leq \tau < 2} \frac{1}{\tau^{1+is_2} (\ln \tau + \ln M_k)} \frac{1}{M_k} \\ &= \frac{1}{M_k^{is_2}} \sum_{1 \leq \tau < 2} \frac{1}{\tau^{1+is_2} (k \ln 2 + \ln M_0 + \ln \tau)} \frac{1}{M_k} \quad (46) \end{aligned}$$

and in the main order we get

$$\frac{1}{k \ln 2} \sum_{1 \leq \tau < 2} \frac{1}{\tau^{1+is}} d\tau.$$

For large k the above Riemann sum can be approximated by the corresponding integral and in the main order we obtain

$$\frac{1}{k \ln 2} \int_1^2 \frac{d\tau}{\tau^{1+is}} + \text{error terms} = \frac{1}{k \ln 2} \frac{1 - 2^{-is_2}}{is_2} + \text{error terms}.$$

The k -th sum (46) is therefore

$$\frac{M_k^{-is_2}}{k \ln 2} \frac{1 - 2^{-is_2}}{is_2} + \text{error terms} = \frac{M_0^{-is_2} 2^{-is_2 k}}{k \ln 2} \frac{1 - 2^{-is_2}}{is_2} + \text{error terms}$$

The series $\sum_k \frac{2^{-is_2k}}{k}$ converges as long as $k \notin \frac{2\pi}{\ln 2}\mathbb{N}$. In fact, since

$$2^{-is_2k} = \sum_{j=0}^k 2^{-is_2j} - \sum_{j=0}^{k-1} 2^{-is_2j} = \frac{1 - 2^{-is_2(k+1)}}{1 - 2^{-is_2}} - \frac{1 - 2^{-is_2k}}{1 - 2^{-is_2}}$$

we can sum by parts and obtain

$$\sum_{k=1}^K \frac{2^{-is_2k}}{k} = - \sum_{k=1}^K \frac{1 - 2^{-is_2k}}{1 - 2^{-is_2}} \frac{1}{k^2} + \text{error terms}$$

which is a convergent series, except for $s \in \frac{2\pi}{\ln 2}\mathbb{N}$. For these values of s we can repeat the construction above using $M'_k = 3^k M'_0$ and obtain a convergent series. \square

The possibility to extend the function $\kappa(s) = \lim_{m \rightarrow \infty} \sum_{n \in \Omega_m} \frac{\mu(n)}{n^s}$ to the regions $\{s \in \mathbb{C} : 0 < \Re s < 1, \Im s \neq 0\}$ is therefore connected with the analytic continuation of the Prime Zeta Function $\sum_p \frac{1}{p^s}$. This function is connected to the Riemann ζ -function via

$$\zeta_P(s) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \ln(\zeta(ks)),$$

see [13]. This formula shows how zeros of the Riemann ζ -function correspond to singularities of ζ_P , clustering along the imaginary axis. For this reason the analytic continuation of ζ_P is only defined in the region $\Re s > 0$. This analytic continuation also has logarithmic singularities on the real line at the points of the form $s = \frac{1}{q}$, where q is square-free. At these points, by (45), $\zeta_P(s)$ and $\kappa(s)$ are consistent. However, for general real $s \in (0, 1)$ the analytic continuation of ζ_P is not compatible with the fact that $\kappa(s) = 0$. This means that the function $\kappa(s)$ for $0 < \Re s < 1$ does not agree with its analytical continuation.

The above argument leaves the following question open: is it true that

- the analytic continuation of $\zeta(s)$,
- the function $\gamma(s)$ defined in (43),
- the function $\frac{1}{\kappa(s)}$, where $\kappa(s)$ is defined in (40)

agree on the vertical line $\Re s = 1$? The answer to this question will cast a new light on the delicate connections between the number-theoretical nature of $\zeta(s)$, $\zeta_P(s)$, and the structure of the ensembles Ω_m we discussed in the previous section.

Acknowledgments We would like to thank the following people for their invaluable comments and remarks: Giovanni Gallavotti, Andrew Granville, Nicholas Katz Alex Kontorovich, Krishnan Mody, Peter Sarnak, Christopher Skinner, Domokos Szász, Anatoly M. Vershik, Ilya Vinogradov. The second author also acknowledges the financial support from NSF, Grant DMS-0600996.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Boca, F.P., Cobeli, C., Zaharescu, A.: Distribution of lattice points visible from the origin. *Commun. Math. Phys.* **213**(2), 433–470 (2000)
2. Bourgain, J.: Correlation bounds for Moebius and Walsh (in preparation)
3. Bourgain, J.: Möbius Schrödinger (to appear)
4. Cellarosi, F., Sinai, Ya.G.: Non-standard limit theorems in Number theory. *Prokhorov Festschrift* (to appear)
5. Davenport, H.: On some infinite series involving arithmetical functions (ii). *Q. J. Math.* **8**, 313–320 (1937)
6. de Bruijn, N.G.: The asymptotic behaviour of a function occurring in the theory of primes. *J. Indian Math. Soc. (N.S.)* **15**, 25–32 (1951)
7. de Bruijn, N.G.: On the number of positive integers $\leq x$ and free of prime factors $> y$. *Nederl. Acad. Wetensch. Proc. Ser. A* **54**, 50–60 (1951)
8. de Bruijn, N.G.: On the number of positive integers $\leq x$ and free prime factors $> y$. II. *Nederl. Akad. Wetensch. Proc. Ser. A* 69 *Indag. Math.* **28**, 239–247 (1966)
9. Dickman, K.: On the frequency of numbers containing primes of a certain relative magnitude. *Ark. Mat. Astr. Fys.* **22** **22A**, 1–14 (1930)
10. Elkies, N.D., McMullen, C.T.: Gaps in $\sqrt{n} \bmod 1$ and ergodic theory. *Duke Math. J.* **123**(1), 95–139 (2004)
11. Erdős, P.: *Wiskundige Opgaven met de Oplossingen*, 21. Problem and Solution Nr. 136 (1963)
12. Erdős, P., Kac, M.: The Gaussian law of errors in the theory of additive number theoretic functions. *Am. J. Math.* **62**, 738–742 (1940)
13. Fröberg, C.-E.: On the prime zeta function. *BIT* **8**(2), 187–202 (1968)
14. Gnedenko, B.V., Kolmogorov, A.N.: Limit distributions for sums of independent random variables. Translated from the Russian, annotated, and revised by K.L. Chung. With appendices by J.L. Doob and P.L. Hsu. Revised edition. Addison-Wesley Publishing Co., Reading, London (1968)
15. Goncharov, V.L.: From the realm of combinatorics (in Russian). *Izv. Akad. Nauk SSSR, Ser. Matem.*, **8** **1**, 3–48 (1944)
16. Granville, A.: Smooth numbers: computational number theory and beyond. In: *Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography*. *Math. Sci. Res. Inst. Publ.*, vol. 44, pp. 267–323. Cambridge University Press, Cambridge (2008)
17. Granville, A., Pomerance, C.: On the least prime in certain arithmetic progressions. *J. Lond. Math. Soc. (2)* **41**(2), 193–200 (1990)
18. Green, B.: On (not) computing the Möbius functions using bounded depth circuits (submitted)
19. Green, B., Tao, T.: The mobius function is strongly orthogonal to nilsequences. *Ann. Math.* (to appear)
20. Heath-Brown, D.R.: Zero-free regions for Dirichlet L -functions, and the least prime in an arithmetic progression. *Proc. Lond. Math. Soc. (3)* **64**(2), 265–338 (1992)
21. Hensley, D.: The convolution powers of the Dickman function. *J. Lond. Math. Soc. (2)* **33**(3), 395–406 (1986)
22. Hildebrand, A.: Integers free of large prime factors and the Riemann hypothesis. *Mathematika* **31**(2), 258–271 (1985)
23. Iwaniec, H., Kowalski, E.: *Analytic number theory*. American Mathematical Society Colloquium Publications, vol. 53. American Mathematical Society, Providence (2004)
24. Kolmogorov, A.N.: Sulla forma generale di un processo stocastico omogeneo. *Atti Acad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat.* **15**(6):805–808; 866–869 (1932)
25. Linnik, U.V.: On the least prime in an arithmetic progression. I. The basic theorem. *Rec. Math. [Mat. Sbornik] N.S.* **15**(57), 139–178 (1944)
26. Linnik, U.V.: On the least prime in an arithmetic progression. II. The Deuring-Heilbronn phenomenon. *Rec. Math. [Mat. Sbornik] N.S.* **15**(57), 347–368 (1944)
27. Liu, J., Sarnak, P.: Integral points on quadrics in three variables whose coordinates have few prime factors (in preparation)
28. Mertens, F.: Ein beitrage zur analytischen zahlentheorie. Ueber die vertheilung der primzahlen. *J. Reine Angew. Math.* **78**, 46–62 (1874)
29. Pappalardi, F.: A survey on k -freeness. In: *Number Theory*. *Ramanujan Math. Soc. Lect. Notes Ser.*, vol. 1, pp. 71–88. Ramanujan Math. Soc., Mysore (2005)

30. Pomerance, C.: A note on the least prime in an arithmetic progression. *J. Number Theory* **12**(2), 218–223 (1980)
31. Sarnak, P.: Möbius randomness and dynamics. Lecture Slides Summer 2010. <http://www.math.princeton.edu/sarnak/Mobius%20lectures%20Summer%202010.pdf>
32. Sarnak, P., Ubis, A.: The horocycle flow at prime values (in preparation)
33. Vershik, A.M.: Asymptotic distribution of decompositions of natural numbers into prime divisors. *Dokl. Akad. Nauk SSSR* **289**(2), 269–272 (1986)
34. Vershik, A.M.: Does a Lebesgue measure in an infinite-dimensional space exist? *Tr. Mat. Inst. Steklova* **259**(Anal. i Osob. Ch. 2):256–281 (2007)
35. Vershik, A.M., Schmidt, A.A.: Limit measures that arise in the asymptotic theory of symmetric groups. I. *Teor. Veroyatnost. i Primenen.* **22**(1), 72–88 (1977)
36. Vershik, A.M., Schmidt, A.A.: Limit measures that arise in the asymptotic theory of symmetric groups. II. *Teor. Veroyatnost. i Primenen.* **23**(1), 42–54 (1978)
37. Xylouris, T.: Über die Linniksche Konstante. Diplomarbeit, Universität Bonn. arXiv:0906.2749v1 (2009)